



# Computing & Software



**DØ main executables**  
**Data and MC production**  
**Data-handling and databases**  
**Central and offsite computing**  
**Random thoughts & issues**





# Thanks to



**Laurent Duflot**  
**Event graphics**

**Marco Verzocchi**  
**DØ analysis software**



**Heidi Schellman**  
**Data access**





# Reconstruction

**DØRECO** is at the heart of our software and **PHYSICS** effort

## **p11**

- has been running on the central farm
- *DØANALYZE* crashes about 10% of the files
- first version of thumbnail deployed
- calorimeter-tracker geometry mismatch

## **p12**

- not certified for data production use
- fixed many problems of p11 (geometry, thumbnail,...)
- will go to online for Level-3 and streaming

## **p13**

- next major version for both data/MC farms
- drop *DØANALYZE*, no more ROOT-tuple
- first version of “physics” DST/TMB(+CalDataChunk)
- successfully deployed GTR+HTF as the default algorithm
- stream zero-bias events for MC on the central farm
- **p13.00.00 cut, met the first milestone !**



# Simulation

## GEANT detector simulation

### p12

- many geometry updates (CAL, forward muon, ...)
- physics groups pushing for MC farms

### p13

- dead-channels handled in unpackers  
implemented in MC for SMT, data for CFT
- SMT geometry updated
- “all materials” are accounted for
- tools available to overlay zero-bias events
- still no relative alignment
- calorimeter task force initial recommendations ?

### Future

More task force recommendations (see Marek's talk)

## Parameterized MC simulation (PMCS)

- p13 will be the first version for “physics analyses”
- very useful for MC-intensive analyses
- need users to use it and help tuning the parameters!



# Trigger

## Level 2/Level 3

- **p11 running now**
  - electron tool based on EM fraction (L2/L3), shower shapes (L3)
  - muon tool with pT cut (L2), no calor. and track matching (L2/L3)
  - tau tool (L3) based on isolation, width, profile, no track match
  - jet tool (L2/L3), stand-alone global track filter (L3)
- **p12 (ready to go online)**
  - L2 jet algorithm moves from 3x3 to 5x5 clustering
  - calorimeter RCP-selectable zero-suppression at L3
  - electrons with track-based road algorithm at L3
  - tau with number of tracks cut, MET tool at L3
  - HT available at L2
- **p13**
  - calorimeter unpacker, jet acoplanarity filter, HT filter at L3
  - improved electron, muon filters
  - 3D primary vertex finder

## Trigger simulation (*TRIGSIM*)

- **simulate L1 hardware, wrap L2/L3 online code for simulation**
- **Welcome Angela Bellavance to be managers of *TRIGSIM* along with Dugan O'Neil and thanks to Serban !**



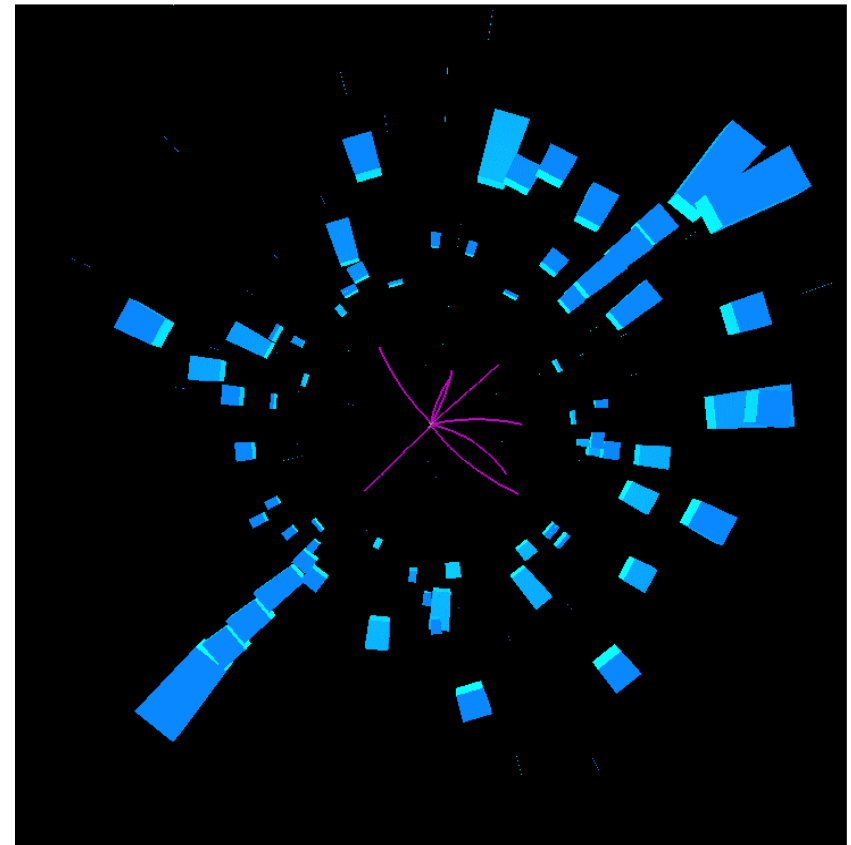


# Event Graphics

## Two packages

- **d0ve**
  - based on X window
  - used extensively online
  - mostly in maintenance mode
- **d0scan**
  - uses Open Inventor as the graphics package and Qt as the base library
  - selectable objects and property views
- **important tool to understand the detector, software and analyses**
- **need users to use them to help the development**

$Z \rightarrow \mu\mu$  candidate  
in calorimeter





# Data Tier

**The central farm currently writes (p11.12.01)**

- **RECO output (~ 500 kB/event) [  $\Rightarrow$  DST ]**
- **ROOT-tuple (~ 115 kB/event)**
- **Thumbnail (~ 30 kB/event)**

**Our goal is to write only**

**DST: 150 kB/event (for data), on tape  
allow limited re-calibration and re-reconstruction**

**TMB: 10 kB/event (for data), on tape and 100% on disks  
high-level information sufficient for most analyses**

**Plan for p13**

- **No more RECO ROOT-tuple**
- **First “physics” version of DST and TMB (+ CalDataChunk)**
- **Switch to TMB-based analysis as soon as possible**

**ROOT format**

- **ROOT-Trees can be built by users from DST/TMB short-term**
- **Saving TMB chunk directly in ROOT in long run?**



# Data Processing

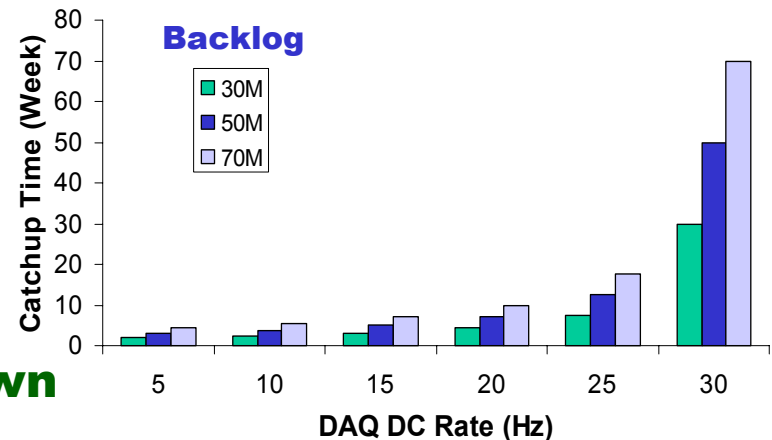
**All initial data processing will be done at Fermilab**

## The “current” farm

- **total CPU cycles ~150 GHz,**
- **most machines are “old”**

## Running RECO p11.12.01

- **processing ~3M events a week,**
- **about 50M backlog since shutdown**

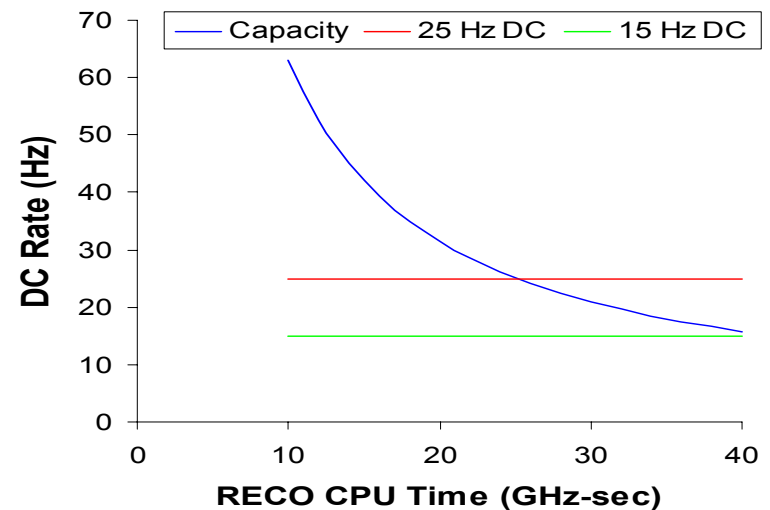


## New farm nodes

- **240 new 1.8 GHz dual nodes**
- **~760 GHz total CPU power**
- **in production end of the month ?**

## Processing issues

- **take 5 weeks to catch up @ 15 Hz**
- **RECO speed < 25 GHz-sec/event**
- **Re-processing needs ?**

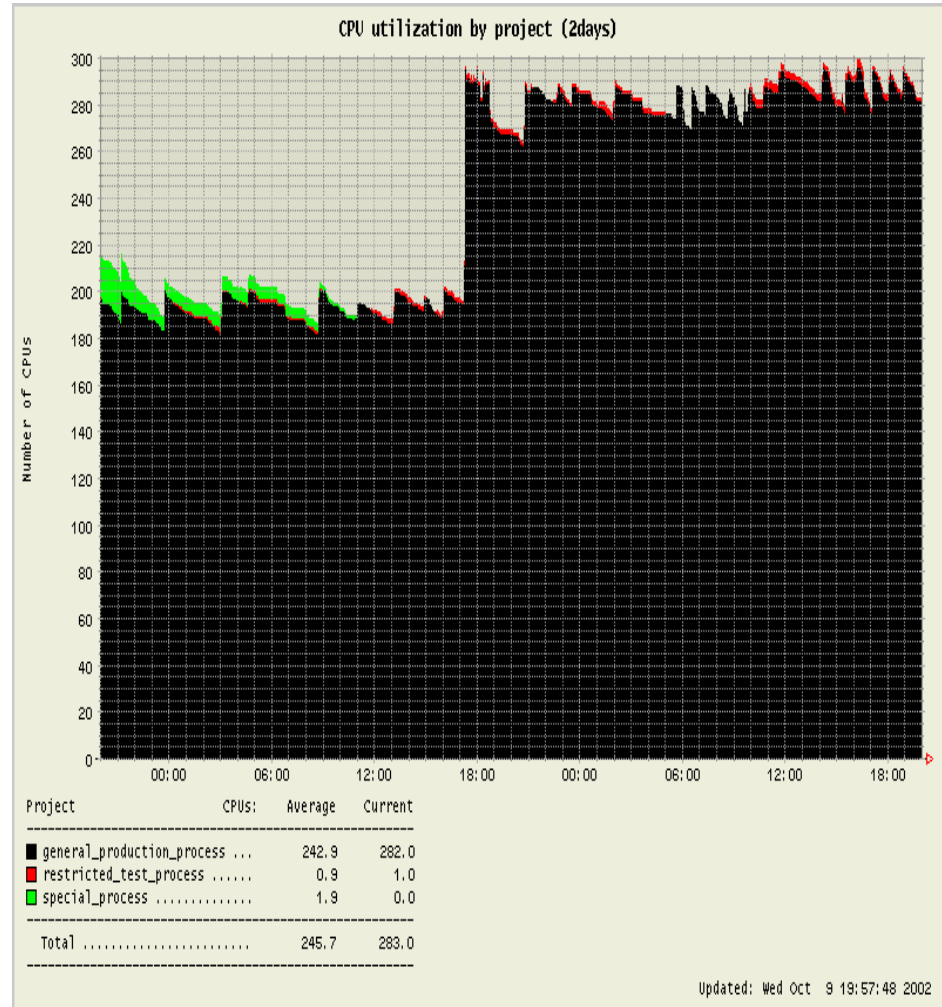






# New Farm Workers

- had rack and power supply problems
- 48 nodes (~150 GHz) in production already
- Capacity ~ 10 Hz DC





# Monte Carlo Production

## Production sites

- all events are officially produced at remote sites:  
Boston, Lancaster, Lyon, NIKHEF, Prague, UTA, ...
- capable of producing ~200k events a day

## Production issues

- running mcp11 now
  - *DØANALYZE* in p11.12.01 won't run on private network  
the problem appears to be fixed in p11.13.00
  - many geometry problems
- p13  $\Rightarrow$  mcp13
  - no *DØANALYZE* by default
  - *TRIGSIM* with a generic trigger list
- migrate to use MC production request system

## Storage proposal (on tape at Fermilab)

- DST with MC Kine and Raw Data Chunks
- TMB with MC Kine Chunk

**Welcome Patrice Lebrun to be coordinators  
along with Iain Bertram of our MC production!**



# SAM and Database

## SAM issues:

- for better or worse, CDF is now a customer
- track our high priority tasks
- tasks addressed in the last two months include
  - adding a file content status word
  - improving Fermilab farm operations
  - resolving Friday melt-down problem on dØmino
  - making ClueDØ SAM operational
  - developing additional monitoring tools
- data handling for remote sites

## Responsibility changes:

- database
  - Taka Yasuda as the new coordinator  
thank Ruth Pordes for her splendid job
- offline shift
  - Kin Yip as the new coordinator for 6 months  
thank Don Coppage for his remarkable effort
  - mostly SAM operational issues now
  - evolving to DØ “helpdesk”



# R2D2 Steering Committee

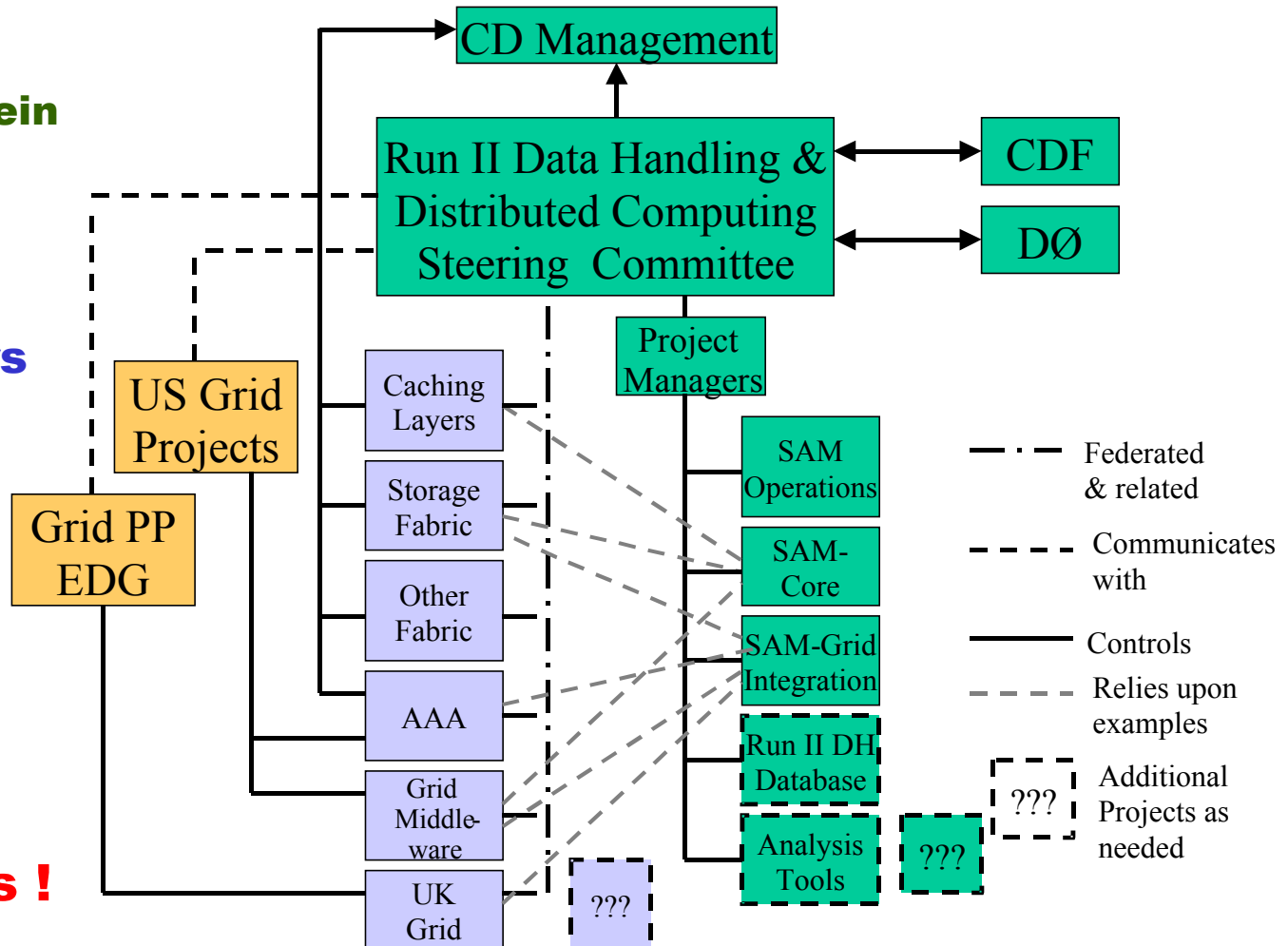
**R2D2: Run 2 Data Handling and Distributed Computing**

## DØ members:

**Amber Boehnlein**  
**Chip Brock**  
**Wyatt Merritt**  
**Jianming Qian**

## Project Managers

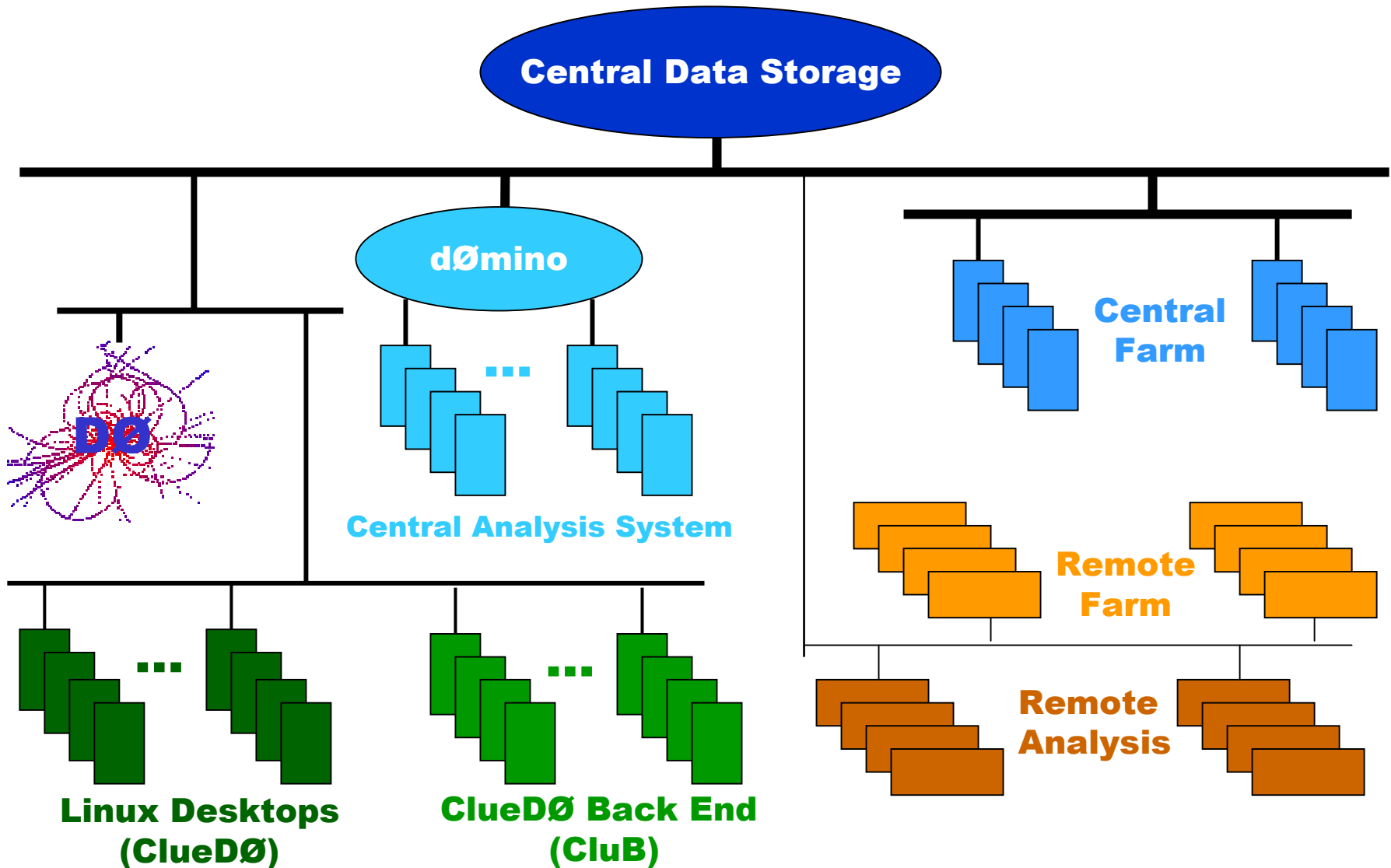
**Lee Lueking**  
**Jeff Tseng**



**Mother of  
all committees !**



# Computing Architecture





# Analysis Computing

## Central Analysis Backend (CAB)

- batch jobs only, using dØmino as I/O
- submit jobs from dØmino (and ClueDØ) using PBS
- development and test on 16 dual 1 GHz nodes
  - 35 GB each SAM and job caches
  - a test version of dØtool installed
- 160 new dual 1.8 GHz nodes being shaken down
  - 500 MB/CPU memory, 80 GB total local cache
- open to users in a few weeks ?
- contributions from institutions are accepted (with strings)

## ClueDØ

- 212 machines from collaborating institutions
- managed by non-CD DØ collaborators
- SAM station in testing, should be available soon

## Disks

- ~ 30 TB on dØmino, half of them for SAM cache
- groups are invited to buy their own disks (prj/tmp)





# Offsite Computing

**Utilizing offsite resources is a part of our computing plan**

## Monte Carlo production

- many sites have been in operation for sometime
- need to develop more sites

## Analysis computing (Nick's talk for details)

- offsite analysis task force (Co-chairs: Hadley, White) established
- **thank Jae Yu for his magnificent job in running DØRACE**
- set up prototype site(s) to demonstrate analyses offsite
  - shipping thumbnails to Karlsruhe
  - the Karlsruhe station pulled over 1 TB from Fermilab in a month
  - need people to use the files at remote sites to help development
- In US
  - UTA has the MRI funding
  - pursuing other funding possibilities

## Reprocessing needs

- improved algorithms, better calibrations, or simply bug fixings
- might not be sufficient resources available at Fermilab
- explore data reprocessing from DST offsite (Michigan, BNL, ... ?)



# Random Remarks

## Online streaming (Adam's talk for details)

- we have to stream to allow flexibility in (re)processing
- some analyses benefit, some don't
- help to design a scenario best for our long-term goal
- important to develop/test tools now

## ROOT-tuple vs Thumbnail

- most analyses are based on ROOT-tuples so far
- ROOT-tuple is a temporary format, not part of our data tier
- ROOT-tuples are large in size, cannot afford them
- p13 is expected to be a lot better than other versions
- p13 produces DST/TMB, not ROOT-tuples
- start using TMB now to help debugging
- plan to make p13 TMB for p11 RECO files



# Immediate Issues

## Streaming

- harness 2-stream test scenario, develop analysis tools
- redesign luminosity database, handle streams
- design a stream scheme for long haul

## p13 (DØRECO, simulation, trigger, graphics)

- available for farms end of the month, aggressive, but doable !?
- functionalities and performances (quality, speed, memory, ...)
- popularize graphics for development, analysis and PR !

## Data/MC production

- commissioning new farm nodes
- catch up with backlog and reprocess with p13
- deploy p13 on MC farms

## Central and offsite computing

- commissioning of CAB, open to users
- ClueDØ SAM station
- analysis, reprocessing at remote sites, data handling issues

## Others

- kcc  $\Rightarrow$  gcc conversion



# User Issues

## High entry barrier

- the software is too complex for most people
- linking takes too long and too much memory
- feeling lost in case of problems

## Documentation

- limited and poor
- the main [c&s web page](#) last modified Sept. 7, 2001
- improvement underway, e.g. [DØ “How To” Documents](#)
- **you can help too !**

## Tutorials

- **thanks to Marco, Laurent and Heidi for their marvelous jobs!**
- tutorials are very time-consuming for experts
- but if effective, will schedule regular topical tutorials
- recruit experienced users to give tutorials

## Operation

- train offline shifters to staff DØ “helpdesk”
- shifters to monitor analysis computing, SAM, ...



# Summary

**Lots of effort and lots of progress!**

**Many issues, challenges remaining.**

**You can help to  
develop  
test  
document !**

**You can make a difference !**